THE DIMENSION AND ENTROPY OF ω-LANGUAGES

YIDING HAO

Abstract. In this paper we will explore the properties of the Hausdoff dimension in Cantor space—the space of strings over arbitrary symbols. It turns out that Hausdorff dimension is closely related to topological entropy, and that the two quantities are equal for closed regular ω-languages. We state and prove those results, and then explore the possibility of extending them to context-free ω-languages and other parts of the Chomsky heirarchy.

Contents

1. Introduction

How complex is a language? Many measures of complexity for languages have been devised. The most famous include *computational complexity*, *Kolmogorov complexity*, and *topological entropy*. It turns out that *Hausdorff dimension*—a measure of the complexity of fractals—can also act as a measure of the complexity of a language. In particular, Hausdorff dimension has an intimate relationship with topological entropy. When applying these measures to sets of infinite strings, we find that Hausdorff dimension is a lower bound on entropy, and under certain conditions, the two measures are equivalent.

In section 5 we will state and prove some important results regarding the equivalency between Hausdorff dimension and entropy. Readers with background in formal language theory, fractal geometry, and information theory may begin with

Date: July 29, 2013.

section 5. Those without background in these areas may find a précis of the prerequisite notions in sections 1–4. In section 6 we will propose a question regarding the relationship between equality of entropy and Hausdorff dimension and the Chomsky heirarchy.

2. Formal Languages

Formal languages are sets of *strings*—sequences of arbitrary symbols. As usual, we refer to a finite set of symbols as an *alphabet* and a set of finite strings over an alphabet as a *language*. For any string w and set W, $|w|$, called the *length* of w, is the number of symbols comprising w, and $|W|$ is the cardinality of W. We denote by ε the *empty string*, a string of length 0. We denote by $\cdot$ *concatenation*, an operation such that $u \cdot v$ is the string consisting of the symbols of u followed by the symbols of v. Finally, for an alphabet Σ, Σ^* is the set of all possible strings of finite length over Σ, and for a language L, L^* is the closure of L under concatenation. From these definitions we clearly see that $(\Sigma^*, \cdot, \varepsilon)$ is the free monoid on Σ.

In addition to the usual definitions, we introduce the following concepts specific to the focus of this paper. First, we define a class of sets that, for reasons mentioned in Section 4, are convenient to work with.

Definition 2.1. Let Σ be an alphabet. Then, let Σ^ω denote the set of all infinite strings over Σ. A subset of Σ^ω is called an ω-language.

Next, we extend the concatenation product so as to allow for alphabets, languages, or ω-languages as arguments.

Definition 2.2. Let $\mathscr{P}(L)$ denote the power set of a set L. Suppose $A \subseteq \Sigma^*$ and $B \in \mathscr{P}(\Sigma^*) \cup \mathscr{P}(\Sigma^\omega)$. Then,

$$A \cdot B := \{a \cdot b | a \in A, b \in B\}.$$

For $u \in \Sigma^*$,

$$u \cdot B := \{u\} \cdot B \quad \text{and} \quad A \cdot u := A \cdot \{u\}.$$

Definition 2.3. Let X be a set. We define $X^0 = \{\varepsilon\}$, and for all $n \in \mathbb{N}$,

$$X^n := \underbrace{X \cdot X \cdot \cdots \cdot X}_{n \text{ times}}.$$

We also want to be able to take infinite concatenation products over a language X. Suppose $\{x_i\}$ is an infinite sequence where for each i, $x_i \in X$. Then, we define

$$x_1 \cdot x_2 \cdot \cdots := \lim_{n\to\infty} x_1 \cdot x_2 \cdot \cdots \cdot x_n$$

and let X^ω denote the ω-language consisting of all infinite strings of the form $x_1 \cdot x_2 \cdot \ldots$, where $x_i \in X$ for every $i \in \mathbb{N}$.

Finally, we introduce the notion of *prefixes*. The notion of prefixation here is somewhat similar to the linguistic notion of prefix. It is important in information theory, and in Section 4 we will use it to induce a topology on $\Sigma^* \cup \Sigma^\omega$.

Definition 2.4. Let u and v be strings over the alphabet Σ. If there exists $w \in \Sigma^*$ such that $u = v \cdot w$, then we say that v is a *prefix* of u, and, adopting the notation from [Str09], we write $v \preceq u$. Moreover, if $w \neq \varepsilon$, then we say that v is a *strict prefix* of u and we write $v \prec u$.

Definition 2.5. A language or ω-language is *prefix-free* if no string in the language is a strict prefix of another string in the language.

Notation 2.6. By $\mathbf{A}(F)$ we denote the set of all prefixes of a set $F \subseteq \Sigma^* \cup \Sigma^\omega$.

With these terminological and notational conventions in place, we will now develop the prerequisite machinery in formal language theory.

2.1. **Formal Grammars.** By *grammar* we mean a formal device for generating formal languages. Formal grammars were originally developed to model the structure of natural languages, and today they are still the foundation of syntactic, semantic, and morphological models in theoretical linguistics. The most well-known type of formal grammar is the *context-free grammar*, whose definition we now recall.

Definition 2.7. A *context-free grammar* is an ordered quadruple (V, Σ, R, S) that satisfies the following.

a. V is an alphabet called the *variables* or *non-terminals.*
b. Σ is an alphabet called the *terminals*, and $V \cap \Sigma = \varnothing$.
c. $R \subseteq V \times (V \cup \Sigma)^*$ is called the *rules*, and its elements are denoted by $A \to B$.
d. $S \in V$ is called the *start symbol.*

Suppose $u, v, w \in (V \cup \Sigma)^*$ and $A \to w \in R$. Then, we say that uAv *yields* uwv, and we write $uAv \Rightarrow uwv$. If $u = v$, $u \Rightarrow v$, or if there exist $u_1, \dots, u_n$ such that

$$u \Rightarrow u_1 \Rightarrow \cdots \Rightarrow u_n \Rightarrow v,$$

then we say that u *derives* v, and we write $u \overset{*}{\Rightarrow} v$. It is through derivations that grammars are able to generate languages.

Definition 2.8. Suppose G is a context-free grammar with start symbol S. The set $\{w | S \overset{*}{\Rightarrow} w\}$ is called the *language generated by* G. A language is called *context-free* if it is generated by some context-free grammar.

One specific type of formal grammar we are interested in is the *regular grammar.*

Definition 2.9. A context-free grammar (V, Σ, R, S) is called a *regular grammar* if all its rules are of the form

$$A \to b \cdot C \quad \text{or } A \to b,$$

where $b \in \Sigma^*$ and $C \in V$. A language is called *regular* if it can be generated by a regular grammar.

It is not difficult to see that the variables, terminals, and rules of a regular grammar correspond to the states, output, and transitions of a finite-state automaton.[1] It is from this correspondence that the terminology for the following definition, which allows us to characterize and generalize the variables in a regular grammar, is motivated.

Definition 2.10. Fix $w \in \Sigma^*$, $W \subseteq \Sigma^* \cup \Sigma^\omega$. The *state of* W *derived by* w is defined as

$$W/w := \{p | w \cdot p \in W\}.$$

[1] Because automata theory is not necessary to prove the main results of this paper, the foundations thereof are not here stated. A clear and comprehensive introductory discussion of the theory may be found in [Sip06].

Suppose W is a regular language generated by the grammar (V, Σ, R, S). If $W/w \neq \varnothing$, then there exists $A \in V$ such that $S \overset{*}{\Rightarrow} w \cdot A$. Because of the equivalence of deterministic and non-deterministic finite state automata, we can assume without loss of generality that A is unique. W/w, then, specifies the language consisting of all strings in Σ^* that are derived by A. Since V must be finite, there can only be finitely many possible values for W/w. This definition allows us to generalize the notion of a regular language and use this generalization to define the class of *regular ω-languages.*

Definition 2.11. If $W \subseteq \Sigma^\omega$ and $|\{W/w | w \in \Sigma^*\}| < \infty$, then we say that W is a *regular ω-language.*

To define the class of *context-free ω-languages* is slightly more complicated. Here we will do so using a theorem about representing regular ω-languages in terms of regular languages.

Theorem 2.12. *An ω-language is regular if and only if it is of the form*

$$\bigcup_{i=1}^{n} U_i \cdot V_i^\omega,$$

where $n \in \mathbb{N}$ and each U_i and V_i is a regular language.

Definition 2.13. An ω-language is *context-free* if it is of the form

$$\bigcup_{i=1}^{n} U_i \cdot V_i^\omega,$$

where $n \in \mathbb{N}$ and each U_i and V_i is a context-free language.

By a theorem analogous to Theorem 2.12, it turns out that this definition is equivalent to the traditional definition of context-free ω-languages as languages accepted by Muller or Büchi automata; the details are given in [Fin08].

3. Fractal Geometry

The word "fractal" itself has no universal definition, but fractals are widely known as sets that display self-similarity. Colloquially, an object is *self-similar* if it appears to be made of copies of itself; more precisely, we may say that a set A in some metric space is self-similar if some subset of A is the image of A under an affine transformation. Famous examples of fractals include the Cantor set, the Sierpinski Triangle, the Koch Snowflake, the Mandelbrot set, and Julia sets. More generally, there are two concepts that we can use to precisely define the notion of a fractal. Firstly, a fractal can be thought of as a set with *fractal dimension* exceeding its topological dimension. Secondly, a fractal can be thought of as the fixed point of a contraction mapping.[2] In this section we develop the theoretical background behind these two definitions.

[2]Because it may describe two different concepts in the literature, we will refrain from using the word "fractal" in a formal context.

3.1. **Hausdorff Dimension.** *Fractal dimensions* are nonnegative numbers (or ∞) associated with subsets of metric spaces. They are used to quantify the subjective notions of "size," "complexity," and "density." More specifically, they describe the relationship between the level of detail with which we can measure the size of an object and the scale from which we view the object.[3] Fractal dimensions also generalize topological dimension and usually coincide with topological dimension for simple geometric shapes in Euclidean space.

There are many different fractal dimensions in fractal geometry. In this paper we are primarily concerned with *Hausdorff–Besicovitch dimension*—also known as *Hausdorff dimension*—and for the remainder of this paper the word "dimension" shall refer to this notion of dimension unless otherwise specified. The definition of Hausdorff dimension relies on *Hausdorff measure*, an outer measure defined on metric spaces.

Definition 3.1. Suppose (X, d) is a metric space. Write

$$\mathbb{L}^\alpha_\delta(S) := \inf_{E_i} \sum_{i=1}^{\infty} [\mathrm{diam}(E_i)]^\alpha,$$

where $\{E_i\}$ is a cover of S and for each i, $\mathrm{diam}(E_i) \leq \delta$. The *$\alpha$-dimensional Hausdorff measure* of $S \subseteq X$ is defined by

$$\mathbb{L}^\alpha(S) := \lim_{\delta \to 0} \mathbb{L}^\alpha_\delta(S).$$

One can easily verify that, for any α, $\mathbb{L}^\alpha$ satisfies nonnegativity, translation invariance, and σ-additivity. In fact, in Euclidean space, Hausdorff measure is linearly related to Lebesgue measure.

Fact 3.2. *Suppose $E \subseteq \mathbb{R}^n$ is a Borel set. Then,*

$$\mathbb{L}^n(E) = \frac{2^n \Gamma(n/2+1)\lambda(E)}{\pi^{n/2}},$$

where λ denotes Lebesgue measure on $\mathbb{R}^n$ and Γ is the Euler gamma function.

In particular, the quantity $\pi^{n/2}/\Gamma(n/2+1)$ is the volume of the unit ball in $\mathbb{R}^n$, and dividing 2^n by this quantity gives us the scaling factor we can use to derive Hausdorff measure from Lebesgue measure. Fact 3.2, therefore, gives us the intuition that Hausdorff measure is simply a generalization of volume. We can now use Hausdorff measure to define Hausdorff dimension.

Definition 3.3. The *Hausdorff dimension* of a subset S of a metric space (X, d) is defined as

$$\dim(S) := \begin{cases} \inf\{\alpha | \mathbb{L}^\alpha(S) = 0\}, & \exists \alpha : \mathbb{L}^\alpha(S) = 0 \\ \infty, & \forall \alpha (\mathbb{L}^\alpha(S) > 0). \end{cases}$$

We again may derive intuition for Hausdorff dimension by thinking of Hausdorff measure as volume. The n-dimensional volume $V_n(S)$ of an m-dimensional figure S is always 0 if $n > m$. If $V_m(S) > 0$, then for $n < m$ there exists an n-dimensional cross-section S' of S such that $V_n(S') > 0$. Therefore, we see that m is the highest dimension in which the volume of S is nonzero—or equivalently, the greatest lower

[3] A rigorous treatment of this subject is given in Chapter V of [Bar93].

bound on the dimension in which the volume of S is 0—and that is exactly how we have defined Hausdorff dimension above.[4]

It should be noted that, unlike topological dimension, the Hausdorff dimension (as well as other fractal dimensions) may assume non-integer values. For example, the Hausdorff dimensions of the Cantor set and the Sierpinski Triangle are $\log_3 2$ and $\log_2 3$, respectively.

3.2. **Iterated Function Systems.** *Iterated Function Systems* (or IFSs) are sets of contraction mappings from a metric space to itself. It turns out that if we define an arbitrary compact subset A of a metric space with certain properties and construct a sequence $\{A_i\}$ such that A_i is the image of A under i iterations of a set of contraction mappings, then $\lim_{i\to\infty} A_i$ does not depend on the choice of the initial set A. In other words, if A and B are arbitrary compact sets and $\{A_i\}$ and $\{B_i\}$ are constructed as described, then it will always hold that

$$\lim_{i\to\infty} A_i = \lim_{i\to\infty} B_i.$$

The set of contraction mappings forms an IFS, and we call $\lim_{i\to\infty} A_i$ the *attractor* of the iterated function system. The attractor of an IFS will, in most cases, exhibit self-similarity.

We will now make the above notions precise and develop the basic theory of attractors of IFSs. Recall that a contraction mapping is defined as follows.

Definition 3.4. Let (X, d) be a metric space, and suppose that $f : X \to X$. If there exists $s \in [0, 1)$ such that, for all $x, y \in X$,

$$d(f(x), f(y)) \leq sd(x, y),$$

then f is called a *contraction mapping on X with contractivity factor s.*

As we have previously mentioned, an IFS is simply a collection of contraction mappings.

Definition 3.5. A *hyperbolic iterated function system*, or *IFS*, is an ordered pair consisting of a complete metric space (X, d) and a finite collection $\{f_i\}$ of functions, where each f_i is a contraction mappings on X with contractivity factor s_i. We denote the IFS by

$$\{X; f_1, \dots, f_n\},$$

where $n = |\{f_i\}|$.

The following well-known theorem, first proven by Stefan Banach, makes it possible for us to obtain many useful and interesting results from contraction mappings. The necessary condition of completeness is the reason why we have required the metric space of an IFS to be complete.

Notation 3.6. Let f be a function. For $n \in \mathbb{N}$, we write

$$f^{\circ n} := \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times}}.$$

[4]It should be noted that, despite the analogy of cross-sectional volume used in the intuitional exposition, if $\alpha < \dim(S)$, then $\mathbb{L}^\alpha(S) = \infty$.

Theorem 3.7 (Contraction Mapping Principle)**.** *If f is a contraction mapping on a complete metric space (X, d), then f has a unique fixed point x, and for all $y \in X$,*

$$\lim_{i\to\infty} f^{\circ i}(y) = x.$$

Because fractals are sets, contraction mappings on points are not interesting in themselves. Instead, we want to consider contraction mappings that are set functions, as the fixed point of such a contraction mapping is a set of points rather than a single point. Furthermore, the self-similar sets of interest in fractal geometry are compact, and therefore the natural metric space on which to consider contraction mappings is the space of compact subsets of some metric space of points. We proceed to introduce notation for the space and then to define the natural distance function induced upon it.

Notation 3.8. Let (X, d) be a metric space. Denote by $\mathscr{H}(X)$ the set of all compact subsets of X.

To define the distance function on $\mathscr{H}(X)$, we recall the usual extension of distance functions on point domains to set domains.

Definition 3.9. Suppose that (X, d) is a complete metric space. For $x \in X$ and $A \subseteq X$, we define

$$d(x, A) := \inf_{y\in A} d(x, y).$$

For $A, B \subseteq X$ and $x \in X$, we define

$$d(A, B) := \sup_{x\in A} d(x, B).$$

Remark 3.10. If A is (sequentially) compact, then $\arg\inf_{y\in A} d(x, y)$ exists and is an element of A. Therefore, when considering sets in $\mathscr{H}(X)$, we can replace "inf" and "sup" in the above definitions with "min" and "max," respectively.

Because we have used the infimum to define the d on $X \times \mathscr{P}(X)$ but the supremum to define d on $\mathscr{P}(X)^2$, d is not symmetrical on $\mathscr{H}(X)^2$, and therefore cannot serve as the distance function on a metric space. We remedy this with the following definition.

Definition 3.11. Let (X, d) be a metric space. Define

$$\begin{aligned} h_d :\ & \mathscr{H}(X)^2 \to [0, \infty) \\ & (A, B) \mapsto \max\{d(A, B), d(B, A)\}. \end{aligned}$$

If (X, d) is complete, then the metric space $(\mathscr{H}(X), h_d)$ is called the *fractal space on X*.

It turns out that contraction mappings on fractal spaces are closely related to contraction mappings on their underlying point spaces. This property allows us to conclude the discussion in this subsection by defining the attractor of an IFS and justifying its existence.

Proposition 3.12. *The following statements hold.*

a. Suppose f is a contraction mapping on (X, d) with contractivity factor s. If, for $A \subseteq X$, we denote $f(A)$ to be the image of A under f, then f is a contraction mapping on $(\mathscr{H}(X), h_d)$ with contractivity factor s.

b. Suppose $\{f_1, \ldots, f_n\}$ *is a collection of contraction mappings on* $(\mathscr{H}(X), h_d)$ *where each* f_i *has contractivity factor* s_i*. For* $A \in \mathscr{H}(X)$*, write*

$$F(A) = \bigcup_{i=1}^{n} f_i(A).$$

Then, F *is a contraction mapping on* $\mathscr{H}(X)$ *with contractivity factor* $\max_i s_i$*.*

Now let (X, d) be a complete metric space, and let $\{X; f_1, \ldots, f_n\}$ be an IFS. If we extend the domain of each f_i in the usual way to $\mathscr{H}(X)$,[5] then by Proposition 3.12(a) each f_i is a contraction mapping on $\mathscr{H}(X)$. If we write $F(A) = \bigcup_{i=1}^{n} f_i(A)$, then by Proposition 3.12(b), F is a contraction mapping on $\mathscr{H}(X)$. The Contraction Mapping Principle then implies that F has a unique fixed point.

Definition 3.13. Let $S = \{X; f_1, \ldots, f_n\}$ be an IFS. For $A \in \mathscr{H}(X)$, write $F(A) = \bigcup_{i=1}^{n} f_i(A)$. The unique fixed point of F is called the *attractor* of S.

4. The Cantor Space

The Cantor set P can be thought of as the set of numbers in $[0, 1]$ whose ternary expansions contain only the digits 0 and 2. By considering the surjective mapping $\zeta : P \to [0, 1]$ that takes the ternary expansion of a number and changes all instances of "2" to "1," one can easily conclude that P is uncountable. If we consider another function that takes a number in $\zeta(P) = [0, 1]$ and maps it to the string of digits in its binary expansion, we obtain $\{\mathtt{0}, \mathtt{1}\}^*$. This set is homeomorphic to the Cantor set, and for that reason we call it the Cantor space. In order to facilitate topological discussion, we will now define a metric on the Cantor space.

Theorem 4.1. *Let* Σ *be an alphabet. Call the function*

$$\begin{aligned} \rho : \quad & (\Sigma^* \cup \Sigma^\omega)^2 \to [0, \infty) \\ & (u, v) \mapsto \begin{cases} |\Sigma|^{-|u \curlywedge v|}, & u \neq v \\ 0, & u = v \end{cases} \end{aligned}$$

the prefix metric. Then, $(\Sigma^* \cup \Sigma^\omega, \rho)$ *is a metric space.*

Proof. We want to verify that ρ is a well-defined metric. It is clear from the definition that $\rho \geq 0$ and that $\rho(u, u) = 0$. It remains to check the triangle inequality.

Fix $u, v, w \in \Sigma^* \cup \Sigma^\omega$. We will now show that

$$\rho(u, v) \leq \max\{\rho(u, w), \rho(v, w)\}, \tag{4.2}$$

which is equivalent to showing that either $(u \curlywedge w) \preceq (u \curlywedge v)$ or $(v \curlywedge w) \preceq (u \curlywedge v)$. This, of course, in turn implies the triangle inequality. Suppose the contrary. Then, $(u \curlywedge v) \prec (u \curlywedge w)$. This means that there exist u' such that $(u \curlywedge w) = (u \curlywedge v) \cdot u'$. By the same argument, there must exist v' such that $(v \curlywedge w) = (u \curlywedge v) \cdot v'$. But since $(u \curlywedge v)$ is a strict prefix of both $(u \curlywedge w)$ and $(v \curlywedge w)$, it must be the case that $u' \curlywedge v' \neq \varepsilon$. This implies that $(u \curlywedge v)$ is not the $\preceq$-maximal prefix of u and v, which is a contradiction. $\square$

Remark 4.3. Inequality (4.2) is called the *strong triangle inequality*, and a metric (such as ρ) that satisfies the strong triangle inequality is called an *ultrametric*.

[5] i.e., by taking $f_i(A)$ to be the image of A under f_i

Observe that the above proof did not depend on whether the arguments of ρ were of finite or infinite length. Therefore, we may consider the subspaces (Σ^*, ρ) and (Σ^ω, ρ) to be metric spaces in their own right. As we will see in the following theorem, if we restrict our metric space to Σ^ω, we derive a useful way to represent open and closed balls.

Theorem 4.4. *Fix $w \in \Sigma^\omega$. Then, $B \subseteq \Sigma^* \cup \Sigma^\omega$ is an open ball around w with radius $r \in (0,1)$ if and only if it is of the form $B = \{x | y \preceq x\}$ for some $y \in (\Sigma^* \cup \Sigma^\omega)$.*

In the literature, open and closed balls are more commonly written in the form $w \cdot \Sigma^\omega$; this form identifies open and closed balls by their common prefix rather than by their radius and center.

Proof of Theorem 4.4. Consider an open ball $B_r(w)$, where $w \neq \varepsilon$. Write

$$n = \left\lceil -\log_{|\Sigma|} r \right\rceil .$$

Then, $\rho(w, w_n) \leq r$ and $\rho(w, w_{n-1}) > r$, so w_n is the smallest prefix of w in B. For any $x \in (\Sigma^* \cup \Sigma^\omega)$, it is obvious that $\rho(w, x) \leq r$ if and only if $w_n \preceq x$, so we can write $B_r(w) = \{x | w_n \preceq x\}$.

Now suppose $B = \{x | y \preceq x\}$ for some y. Fix any $x \in B$, and choose z so that $\rho(x, z) \leq \rho(x, y)$. This implies that

$$y \preceq (x \curlywedge y) \preceq (x \curlywedge z),$$

so it must be the case that $z \in B$, hence $B = B_{\rho(x,y)}(x)$. $\square$

Remark 4.5. It is not difficult to see that the above argument extends Theorem 4.4 to closed balls as well. Thus, all balls in Σ^ω are simultaneously open and closed (or *clopen*).

Because the distance between a string $u \in \Sigma^*$ and any other string can never exceed $|\Sigma|^{-|u|}$, Theorem 4.4 does not hold in Σ^*. Therefore, for the purposes of this paper it is more convenient to work primarily with the space of ω-languages rather than with Σ^*, and the literature in this area is more extensive for ω-languages than for languages in Σ^*.

We now define a few functions and operations in Cantor space that will be useful to our discussion.

Definition 4.6. If $W \subseteq \Sigma^*$, then the *structure function of* W is defined as

$$\begin{aligned} \mathsf{s}_W : \quad & \mathbb{N} \to \mathbb{N} \\ & n \mapsto |W \cap \Sigma^n|. \end{aligned}$$

In other words, $\mathsf{s}_W(n)$ is the number of strings of length n in the language W. If $W \subseteq \Sigma^\omega$, then we define

$$\mathsf{s}_W \equiv \mathsf{s}_{\mathbf{A}(W)}.$$

(Recall that $\mathbf{A}(W)$ is the set of all strict prefixes of W.)

Definition 4.7. Suppose $V \subseteq \Sigma^*$. Then, the δ*-limit* of V is defined as

$$V^\delta := \{v | v \in \Sigma^\omega, |\mathbf{A}(\{v\}) \cap V| \geq \aleph_0\}.$$

We can think of the δ-limit in the following way. Suppose $v \in V^\delta$. Then, we can construct a sequence $\{v_i\} \subseteq V$ such that, for each i, $v \curlywedge v_i \neq \varepsilon$ and $v_i \prec v_{i+1}$. Therefore, $\rho(v, v_i)$ is strictly decreasing by a factor of at least $1/|\Sigma|$, so

$$\lim_{i \to \infty} \rho(v, v_i) = 0.$$

This means that v is a limit point of V in $\Sigma^* \cup \Sigma^\omega$. In fact, it is not difficult to see that V^δ contains all the limit points of V, and therefore $V \cup V^\delta$ is the closure of V, which we denote by $\overline{V}$. Similarly, if V is an ω-language, then $\overline{V} = \mathbf{A}(V)^\delta$.

4.1. **The Hausdorff Measure in Cantor Space.** The property that we are able to write clopen balls in Σ^ω in the form $v \cdot \Sigma^\omega$ allows us to give an alternate definition for Hausdorff measure specific to the Cantor space. Recall the term

$$\mathbb{L}^\alpha_\delta(S) = \inf_{E_i} \sum_{i=1}^{\infty} [\operatorname{diam}(E_i)]^\alpha, \tag{4.8}$$

which converges to $\mathbb{L}^\alpha$ as $\delta \to 0$. $\{E_i\}$ is a covering of S, and we can assume without loss of generality that each E_i is a clopen ball. Thus, to each E_i we can associate a string $v_i \in \Sigma^*$ such that

$$E_i = v_i \cdot \Sigma^\omega.$$

Note that, for any v, $\operatorname{diam}(v \cdot \Sigma^\omega) = |\Sigma|^{-|v|}$. Now let V be the set of all v_is. Then, (4.8) becomes

$$\mathbb{L}^\alpha_\delta(S) = \inf_{V : |v_i| \geq k(\delta)} \left(\sum_{v \in V} |\Sigma|^{-\alpha|v|} \right), \tag{4.9}$$

where $k(\delta) = -\log_{|\Sigma|}(\delta)$ and $S \subseteq \bigcup_{i=1}^n v_i \cdot \Sigma^\omega$. We denote the sum within the parentheses in (4.9) by $\mathbb{L}^\alpha_V(S)$. This gives us

$$\begin{aligned}
\mathbb{L}^\alpha(S) &= \lim_{\delta \to 0} \mathbb{L}^\alpha_\delta(S) \\
&= \lim_{\delta \to 0} \left(\inf_{V : |v_i| \geq k(\delta)} \mathbb{L}^\alpha_V(S) \right) \\
&= \lim_{k \to \infty} \left(\inf_{V : |v_i| \geq k} \mathbb{L}^\alpha_V(S) \right) \\
&= \lim_{k \to \infty} \left(\inf_{V : |v_i| = k} \mathbb{L}^\alpha_V(S) \right). \qquad (4.10)
\end{aligned}$$

Now observe that

$$\arg \inf_{V : |v_i| = k} \mathbb{L}^\alpha_V(S) = \{v_i | S/v_i \neq \varnothing, |v_i| = k\}.$$

Therefore, continuing (4.10), we get

$$\begin{aligned}
\lim_{k \to \infty} \left(\inf_{V : |v_i| = k} \mathbb{L}^\alpha_V(S) \right) &= \lim_{k \to \infty} \left(\inf_{V : |v_i| = k} \sum_{v \in V} |\Sigma|^{-\alpha|v|} \right) \\
&= \lim_{k \to \infty} \mathsf{s}_S(k) |\Sigma|^{-\alpha k}. \qquad (4.11)
\end{aligned}$$

Thus, instead of thinking of balls with arbitrary small diameter, we can instead define Hausdorff measure by thinking of balls with arbitrarily large common prefixes. This will prove to be useful in the ensuing arguments.

5. Entropy and Hausdorff Dimension

In information theory, *entropy* is the length of the smallest string required to convey a certain message. In [Sha48], Claude Shannon considers a device that stochastically outputs strings w from a language $W \subseteq \Sigma^* \cup \Sigma^\omega$ with probability $p(w)$ and defines entropy as

$$H(W) := -\sum_{w \in W} p(w) \log_{|\Sigma|} p(w).$$

The name *entropy* is motivated by the resemblance of this expression to the formula given by Ludwig Boltzmann for calculating the entropy of thermodynamic systems [HPS92].

In this paper, we consider *topological entropy*—a generalization of Shannon entropy that Shannon originally referred to as *channel capacity*.

Definition 5.1. The *topological entropy* of a language $W \subseteq \Sigma^*$ is defined as

$$\mathsf{H}(W) := \limsup_{n \to \infty} \frac{\log_{|\Sigma|}[1 + \mathsf{s}_W(n)]}{n}.$$

If $W \subseteq \Sigma^\omega$, then we define

$$\mathsf{H}(W) \equiv \mathsf{H}[\mathbf{A}(W)].$$

For the rest of this paper, the word *entropy* shall refer to topological entropy.

It turns out that, in Cantor space, the entropy of a language is equivalent to its *upper Minkowski dimension*—another fractal dimension. We denote this fractal dimension by $\overline{\dim}_B$, and we define the *lower Minkowski dimension* as

$$\underline{\dim}_B := \liminf_{n \to \infty} \frac{\log_{|\Sigma|}[1 + \mathsf{s}(n)]}{n}.$$

When the sequence $\log_{|\Sigma|}[1 + \mathsf{s}(n)]/n$ converges, we have $\overline{\dim}_B = \underline{\dim}_B$. In that case, we denote this quantity by $\dim_B$, and we call it the *Minkowski dimension*.[6] In Euclidean space, the Minkowski dimension is equivalent to the Hausdorff dimension for sets that satisfy a condition known as the *open set condition* (or *OSC*). Therefore, to relate entropy to Hausdorff dimension, and in particular to determine conditions under which the two quantities are equal, is a natural line of inquiry in this subject area.

Immediately, we can see that the lower Minkowski dimension (and therefore the entropy) is an upper bound on the Hausdorff dimension.

Proposition 5.2 ([Sta89])**.** *For all* $F \subseteq \Sigma^\omega$,

$$\dim(F) \leq \underline{\dim}_B(F) \leq \overline{\dim}_B(F) = \mathsf{H}(F).$$

Proof. The second inequality is obvious; we proceed to prove the first. Suppose that $\alpha < \dim(F)$, so that $\mathbb{L}^\alpha(F) > 0$. Then, by formula (4.11), for some constant $\varepsilon > 0$,

$$\mathbb{L}^\alpha(F) = \lim_{n \to \infty} \frac{\mathsf{s}_F(n)}{|\Sigma|^{\alpha n}} > \varepsilon.$$

[6]This fractal dimension is also known as *Minkowski–Bouligand dimension* or *box-counting dimension*.

Therefore, for all n large enough, we have

$$\frac{\mathsf{s}_F(n)}{|\Sigma|^{\alpha n}} > \varepsilon,$$

hence

$$\log_{|\Sigma|} \mathsf{s}_F(n) > \alpha n + \log_{|\Sigma|} \varepsilon.$$

Superadditivity of the limit inferior then gives us

$$\underline{\dim}_B(F) > \alpha + \liminf_{n\to\infty} \frac{1 + \log_{|\Sigma|} \varepsilon}{n} = \alpha.$$

Since the above result holds for all $\alpha < \dim(F)$, the inequality must hold. □

The next theorem, which requires some lemmas, gives a more explicit relationship between entropy and dimension.

Theorem 5.3 ([Sta93])**.** *If $F \in \Sigma^\omega$, then*

$$\dim(F) = \inf_{W : F \subseteq W^\delta} \mathsf{H}(W).$$

Definition 5.4. Let $W \subseteq \Sigma^* \cup \Sigma^\omega$. The *structure-generating function of W* is defined as

$$\begin{aligned} \mathfrak{s}_W : \quad & [0, \infty) \to [0, \infty] \\ & t \mapsto \sum_{i=1}^{\infty} \mathsf{s}_W(i) t^i. \end{aligned}$$

Remark 5.5. Notice that, for any $F \subseteq \Sigma^\omega$ and $V \subseteq \Sigma^*$ such that $F \subseteq V \cdot \Sigma^\omega$,

$$\mathbb{L}^\alpha_V(F) = \sum_{v \in V} |\Sigma|^{-\alpha|v|} = \sum_{i=1}^{\infty} \mathsf{s}_V(i) |\Sigma|^{-\alpha} = \mathfrak{s}_V(|\Sigma|^{-\alpha}).$$

This allows us to generalize formula (4.11) to arbitrary sets V defining clopen covers of an ω-language.

Remark 5.6. We know that $\mathfrak{s}_W$ is a power series with radius of convergence

$$\mathsf{rad}(W) := \liminf_{n\to\infty} \frac{1}{\sqrt[n]{\mathsf{s}_W(n)}}.$$

Hence, if $\mathfrak{s}_W(|\Sigma|^{-\alpha}) < \infty$, then

$$|\Sigma|^{-\alpha} \leq \liminf_{n\to\infty} \frac{1}{\sqrt[n]{\mathsf{s}_W(n)}},$$

which implies that

$$\alpha \geq \limsup_{n\to\infty} \frac{\log_{|\Sigma|} \mathsf{s}_W(n)}{n} = \mathsf{H}(W).$$

By reversing this argument, we can show that the opposite direction holds as well.

With these remarks in mind, we state the following lemma.

Lemma 5.7. *Suppose that $F \subseteq \Sigma^\omega$. If $\mathbb{L}^\alpha(F) = 0$, then there exists $W \subseteq \Sigma^*$ such that $F \subseteq W^\delta$ and $\mathsf{s}_W(|\Sigma|^{-\alpha}) < \infty$.*

Proof. For each $i \in \mathbb{N}$, let $W_i \subseteq \Sigma^*$ satisfy $F \subseteq W_i \cdot \Sigma^\omega$ and

$$\mathbb{L}^\alpha_{W_i} = \sum_{w \in W_i} \frac{1}{|\Sigma|^{\alpha|w|}} < \frac{1}{|\Sigma|^i}.$$

This implies that the length of each string in W_i is strictly greater than $i\alpha^{-1}$. Now write $W = \bigcup_{i=1}^\infty W_i$. Then,

$$\mathfrak{s}_W\left(\frac{1}{|\Sigma|^\alpha}\right) = \sum_{i=1}^\infty \frac{\mathsf{s}_W(i)}{|\Sigma|^{\alpha i}}.$$

Without loss of generality, assume that for all i, each string in W_i has the same length, and that $\{W_i\}$ is strictly increasing in the length of strings in W_i. Then, formula (4.11) gives us

$$\sum_{i=1}^\infty \frac{\mathsf{s}_W(i)}{|\Sigma|^{\alpha i}} \leq \sum_{i=1}^\infty \mathbb{L}^\alpha_{W_i}(F) < \sum_{i=1}^\infty \frac{1}{|\Sigma|^i} < \infty.$$

It remains to show that $F \in W^\delta$. Since every W_i satisfies $F \subseteq W_i \cdot \Sigma^\omega$, F must have infinitely many prefixes in W. This completes the proof. $\square$

The next lemma implies the converse of the previous one.

Lemma 5.8. *If* $\mathfrak{s}_V(|\Sigma|^{-\alpha}) < \infty$, *then* $\mathbb{L}^\alpha(V^\delta) = 0$.

Proof. For any $i \in \mathbb{N}$, write

$$V \supseteq V_i := \{v || \mathbf{A}(v) \cap V| = i + 1\}.$$

Then, it is clear that $|v| > i$ for all $v \in V_i$ and $V^\delta \subseteq V_i \cdot X^\omega$ for all i, so

$$\begin{aligned} \mathbb{L}^\alpha(V^\delta) &= \lim_{i\to\infty} \mathbb{L}^\alpha_{V_i}(V^\delta) \\ &= \lim_{i\to\infty} \mathfrak{s}_{V_i}(|\Sigma|^{-\alpha}). \end{aligned}$$

Observe that

$$\sum_{i=1}^\infty \mathfrak{s}_{V_i}(|\Sigma|^{-\alpha}) = \sum_{i=1}^\infty \sum_{j=1}^\infty \mathsf{s}_{V_i}(j)(|\Sigma|^{-\alpha j}).$$

It follows immediately from the hypotheses that $\mathfrak{s}_{V_i}(|\Sigma|^{-\alpha}) \leq \mathfrak{s}_V(|\Sigma|^{-\alpha}) < \infty$, so $\mathfrak{s}_{V_i}(|\Sigma|^{-\alpha})$ converges absolutely. Therefore,

$$\sum_{i=1}^\infty \sum_{j=1}^\infty \mathsf{s}_{V_i}(j)(|\Sigma|^{-\alpha j}) = \sum_{j=1}^\infty \sum_{i=1}^\infty \mathsf{s}_{V_i}(j)(|\Sigma|^{-\alpha j}).$$

Since $V = \bigcup_{i=1}^\infty V_i$ and all V_is are disjoint,

$$\begin{aligned} \sum_{j=1}^\infty \sum_{i=1}^\infty \mathsf{s}_{V_i}(j)(|\Sigma|^{-\alpha j}) &= \sum_{j=1}^\infty \mathsf{s}_V(j)(|\Sigma|^{-\alpha j}) \\ &= \mathfrak{s}_V(|\Sigma|^{-\alpha}) \\ &< \infty. \end{aligned}$$

Consequently,

$$\lim_{i\to\infty} \mathfrak{s}_{V_i}(|\Sigma|^{-\alpha}) = 0,$$

which completes the proof. $\square$

Because $A \subseteq B$ clearly implies $\mathbb{L}^\alpha(A) \leq \mathbb{L}^\alpha(B)$, we can now combine Lemmas 5.7 and 5.8 into a bidirectional statement.

Corollary 5.9. *Suppose that $F \subseteq \Sigma^\omega$. Then, $\mathbb{L}^\alpha(F) = 0$ if and only if there exists $W \subseteq \Sigma^*$ such that $F \subseteq W^\delta$ and $\mathsf{s}_W(|\Sigma|^{-\alpha}) < \infty$.*

The proof of Theorem 5.3 is now easy.

Proof of Theorem 5.3. By definition,

$$\begin{aligned}
\dim(F) &= \inf\{\alpha | \mathbb{L}^\alpha(F) = 0\} \\
&= \inf\{\alpha | \exists W \subseteq \Sigma^*[F \subseteq W^\delta, \mathsf{s}_W(|\Sigma|^{-\alpha}) < \infty]\} \ \because \text{Corollary 5.9} \\
&= \inf\{\alpha | \exists W \subseteq \Sigma^*[F \subseteq W^\delta, \mathsf{H}(W) \leq \alpha]\} \ \because \text{Remark 5.5} \\
&= \inf_{W : F \subseteq W^\delta} \mathsf{H}(W).
\end{aligned}$$

□

5.1. **Conditions for Equality.** We have now established entropy as an upper bound for Hausdorff dimension and proven an explicit relationship between the two quantities. However, it was earlier mentioned that a more interesting question is that of finding conditions under which equality between entropy and dimension holds. This subsection is dedicated to finding such conditions.

One such condition is related to the growth of an ω-language F; i.e., the asymptotic behavior of s_F in relation to $\mathsf{s}_{F/w}$ for arbitrary $w \in \Sigma^*$.

Definition 5.10. We say that an ω-language F has *uniformly bounded growth* if there exists a function $g : \mathbb{N} \to \mathbb{N}$ such that

$$\log_{|\Sigma|} g(n) = o(n)$$

and

$$\mathsf{s}_{F/w} \leq g(|w|) \cdot \mathsf{s}_F(n)$$

for all $w \in \Sigma^*$ and $n \in \mathbb{N}$.

Theorem 5.11 ([Sta89])**.** *If $F \subseteq \Sigma^\omega$ is closed and has uniformly bounded growth, then*

$$\dim(F) = \mathsf{H}(F).$$

This theorem follows from the following property of closed ω-languages.

Lemma 5.12. *Suppose F is a closed ω-language. Then, for every $\alpha < \mathsf{H}(F)$ and $\varepsilon \in (0, \mathsf{H}(F) - \alpha]$, there exist infinitely many $n \in \mathbb{N}$ satisfying*

$$\sum_{v \in V} \frac{\mathsf{s}_F(n - |v|) \cdot |\Sigma|^{\varepsilon |v|}}{\mathsf{s}_F(n)} \leq \mathbb{L}^\alpha_V(F)$$

for all V such that $F \subseteq V \cdot \Sigma^\omega$, where $\mathsf{s}_F \equiv 0$ on $\mathbb{Z} \cap (-\infty, 0)$.

Proof. Since $\alpha + \varepsilon < \mathsf{H}(F)$, by Remark 5.6,

$$|\Sigma|^{-(\alpha+\varepsilon)} \geq \mathsf{rad}(W) = \liminf_{n \to \infty} \mathsf{s}_F(n)^{-1/n}.$$

This implies

$$1 \leq \limsup_{n \to \infty} \mathsf{s}_F(n)^{1/n} |\Sigma|^{-(\alpha+\varepsilon)},$$

so $\mathsf{s}_F(n)|\Sigma|^{-n(\alpha+\varepsilon)}$ must be unbounded above for n large. Consequently, there exist infinitely many n that satisfy

$$\mathsf{s}_F(n)|\Sigma|^{-n(\alpha+\varepsilon)} \geq \mathsf{s}_F(i)|\Sigma|^{-i(\alpha+\varepsilon)}$$

for all $i < n$. Now choose $i = n - |v|$ for arbitrary $v \in V$. We get

$$\mathsf{s}_F(n)|\Sigma|^{-n(\alpha+\varepsilon)} \geq \mathsf{s}_F(n-|v|)|\Sigma|^{|v|(\alpha+\varepsilon)-n(\alpha+\varepsilon)},$$

hence

$$\sum_{v\in V} \frac{\mathsf{s}_F(n-|v|)|\Sigma|^{\varepsilon|v|}}{\mathsf{s}_F(n)} \leq \sum_{v\in V} |\Sigma|^{-\alpha|v|} = \mathbb{L}_V^\alpha(F).$$

□

Proof of Theorem 5.11. Let V be a language that satisfies the following:

a. V is prefix-free;
b. for every $\phi \in F$, there exists $v \in V$ such that $v \preceq \phi$; and
c. for every $v \in V$, there exists $\phi \in F$ such that $v \prec \phi$.

Then, it is clear that $F \subseteq V \cdot \Sigma^\omega$, so fixing $\alpha \in (0, \mathsf{H}(F))$ and $\varepsilon \in (0, \mathsf{H}(F) - \alpha]$, Lemma 5.12 gives us

$$\sum_{v\in V} \frac{\mathsf{s}_F(n-|v|)|\Sigma|^{\varepsilon|v|}}{\mathsf{s}_F(n)} \leq \mathbb{L}_V^\alpha(F) \tag{5.13}$$

for infinitely many $n \in \mathbb{N}$. For n large enough,

$$\mathsf{s}_F(n) = \sum_{v\in V} \mathsf{s}_{F/v}(n-|v|). \tag{5.14}$$

Since F has uniformly bounded growth,

$$\mathsf{s}_{F/v}(n-|v|) \leq g(|v|)\mathsf{s}_F(n-|v|), \tag{5.15}$$

where g satisfies $\log_{|\Sigma|} g(n) = o(n)$; thus, $g(n) \leq |\Sigma|^{\varepsilon n}$ for n large enough. Combining (5.13), (5.14), and (5.15) yields

$$\begin{aligned} \mathsf{s}_F(n) &= \sum_{v\in V} \mathsf{s}_{F/v}(n-|v|) \\ &\leq \sum_{v\in V} \mathsf{s}_F(n-|v|)|\Sigma|^{\varepsilon|v|} \\ &\leq \mathbb{L}_V^\alpha(F)\mathsf{s}_F(n). \end{aligned}$$

Hence, $\mathbb{L}_V^\alpha(F) \geq 1$, so $\alpha \leq \dim(F)$. □

6. Regular and Context-Free Languages

One of the interesting results from [Sta89] is the following.

Lemma 6.1. *All regular ω-languages have uniformly bounded growth.*

Of course, this lemma leads to an obvious corollary.

Corollary 6.2. *If an ω-language F is regular and closed, then*

$$\dim(F) = \mathsf{H}(F).$$

This result motivates the following question: what is the relationship between the place of an ω-language within the Chomsky heirarchy and the conditions under which dimension and entropy are equivalent? It is with the proposal of this question that we will conclude our discussion. In this section, we first provide a proof for Lemma 6.1, and then we will propose two methods for answering this question for context-free languages and survey some relevant results.

Proof of Lemma 6.1. We want to show that, for all regular sets $F \subseteq \Sigma^\omega$, strings $w \in \Sigma^*$, and numbers $n \in \mathbb{N}$, there exists a function g such that

$$\mathsf{s}_{F/w}(n) \leq g(|w|)\mathsf{s}_F(n)$$

and $\log_{|\Sigma|} g(|w|) = o(|w|)$. Since it is clear that, for any n, $\mathsf{s}_{F/w}(n)$ can never exceed $|\Sigma|^{|w|}\mathsf{s}_F(n)$, for each w there must exist a constant C_w such that

$$\mathsf{s}_{F/w}(n) \leq C_w \mathsf{s}_F(n).$$

Now simply choose $g \equiv \max_w C_w$, and the lemma follows. $\square$

6.1. **Growth Rates.** One way to approach this question is to try to find conditions under which a context-free ω-language has subexponential growth. A considerable amount of work has already been done on the growth rate of context-free languages. [BG02] and [Inc01] independently proved the following result.

Definition 6.3. A language is *bounded* if it is the subset of $w_1^* \cdot w_2^* \cdot \ldots w_n^*$ for some finite language $\{w_1, \ldots, w_n\}$.

Theorem 6.4 ([Inc01])**.** *If L is a context-free language, then either $\mathsf{s}_L(n) = O(k^n)$ for some $k > 1$ or $\mathsf{s}_L(n) = O(n^k)$ for some $k \geq 0$.*

Theorem 6.5 ([BG02])**.** *If L is a context-free language, then either $\mathsf{s}_L(n) = O(k^n)$ or L is bounded.*

This allows us to quickly specify whether the growth of a context-free language is exponential or polynomial using only a small amount of information about the structure of a language. If, however, the grammar of the language is known, [GLS05] provides a method for computing the growth rate of the language from the grammar.

The results mentioned above only apply to subsets of Σ^*; however, we are interested in subsets of Σ^ω. In order to apply them to ω-languages, we recall that, for any $F \subseteq \Sigma^\omega$, s_F and $\mathsf{H}(F)$ are defined in terms of $\mathbf{A}(F)$. We can then use the following result.

Lemma 6.6 (Init-lemma, [Sta97])**.** *If F is a context-free ω-language, then $\mathbf{A}(F)$ is a context-free language.*

With these tools at hand, the final step would be to relate the growth rate of F/w to that of F. One possible way to do so is to consider F/w to be a set of *suffixes*.

Definition 6.7. Let $x, y \in \Sigma^*$. We say that x is a *suffix* of y if there exists $z \in \Sigma^*$ such that $z \cdot x = y$.

For $L \subseteq \Sigma^*$, the language $\mathsf{suffix}(L) := \{L/w | w \in \Sigma^*\}$ is simply the set of all suffixes of strings in L. There is a well-known algorithm with which on can, given a context-free grammar that generates a language L, derive another context-free grammar that generates $\mathsf{suffix}(L)$. Thus, using these tools, one may potentially find a relationship between the growth rates of $\mathbf{A}(F)$ and $\mathsf{suffix}(\mathbf{A}(F))$ for an ω-language F, and use this relationship to find conditions under which F has uniformly bounded growth.

6.2. **The Entropy of ω-Powers.** An *ω-power language* is an ω-language of the form W^ω, where $W \subseteq \Sigma^*$. They make an appearance in the following result.

Theorem 6.8 ([Sta05])**.** *Let $W \subseteq \Sigma^*$. Then,*

$$\dim(W^\omega) = -\log_{|\Sigma|} \mathsf{rad}(W^*) = \mathsf{H}(W^*).$$

To see why this theorem is useful, recall that every context-free ω-language can be represented in the form $\bigcup_i U_i \cdot V_i^\omega$, where each U_i and V_i is a context-free language. For any context-free language W, [Kui70] gives a method for calculating $\mathsf{H}(W^*)$. Thus, all that is needed is information on $\mathsf{H}(W^\omega)$; if $\mathsf{H}(W^\omega) = \mathsf{H}(W^*)$, then $\dim(W^\omega) = \mathsf{H}(W^\omega)$.

Acknowledgments. I would like to thank my mentor, Jonathan Stephenson, for the guidance he has provided throughout this REU. I would also like to thank Prof. J. Peter May and the Department of Mathematics for giving me the opportunity to participate in this program. This summer I have learned a great deal of and about mathematics, and I am truly grateful to these individuals for making this experience possible.

References

[Bar93] M. Barnsley. *Fractals Everywhere*. Academic Press Professional, 2^{nd} edition, 1993.

[BG02] M. R. Bridson and R. H. Gilman. Context-free languages of sub-exponential growth. *Journal of Computer and System Sciences*, 64:308–310, 2002.

[Fin08] O. Finkel. Topological complexity of context-free ω-languages: A survey. *CoRR*, abs/0806.1413, 2008.

[GLS05] H. Gruber, J. Lee, and J. Shallit. Enumerating regular expressions and their languages. In M. Domaratzki, A. Okhotin, K. Salomaa, and S. Yu, editors, *Implementation and Application of Automata*, volume 3317 of *Lecture Notes in Computer Science*, pages 2–22. Springer-Verlag, 2005.

[HPS92] G. Hansel, D. Perrin, and I. Simon. Compression and entropy. In *In Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, pages 515–528. Springer-Verlag, 1992.

[Inc01] R. Incitti. The growth function of context-free languages. *Theoretical Computer Science*, 255:601–605, 2001.

[Kui70] W. Kuich. On the entropy of context-free languages. *Information and Control*, 16:173–200, 1970.

[Sha48] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, 1948.

[Sip06] M. Sipser. *Introduction to the Theory of Computation*. Thomson Course Technology, 2^{nd} edition, 2006.

[Sta89] L. Staiger. Combinatorial properties of the hausdorff dimension. *Journal of Statistical Planning and Inference*, 23:95–100, 1989.

[Sta93] L. Staiger. Kolmogorov complexity and hausdorff dimension. In *Fundamentals of Computation Theory*, pages 434–443. Springer-Verlag, 1993.

[Sta97] L. Staiger. ω-languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3 of *Lecture Notes in Computer Science*, pages 339–387. Springer-Verlag, 1997.

[Sta05] L. Staiger. Infinite iterated function systems in cantor space and the hausdorff measure of ω-power languages. *International Journal of Foundations of Computer Science*, 16:787–802, 2005.

[Str09] B. Stratford. The topology of language. *Journal of Mathematical Psychology*, 53:502–509, 2009.